International Journal of Computers and Informatics

Journal Homepage: https://www.ijci.zu.edu.eg

Int. j. Comp. Info. Vol. 4 (2024) 53–67

Paper Type: Review Article

Systematic Analysis of EEG Emotion Recognition: Subject-Dependent vs. Independent Approaches

Ahmed Elrefaiy [1,*]

[1] Department of Computer Science, Faculty of Computers and Informatics, Zagazig University, 44519, Egypt; AMNasser@zu.edu.eg.

Received: 12 May 2024 **Revised:** 10 Aug 2024 **Accepted:** 07 Sep 2024 **Published:** 10 Sep 2024

Abstract

Electroencephalography (EEG) signals, which show human brain activity, hold potential beyond medical diagnosis. They are particularly valuable for emotion recognition, while analyzing these EEG signal to distinguish between various emotional states. In this area, research focuses on two tasks including subject-dependent and subject-independent tasks. While machine learning and deep learning models have been developed to use EEG data for recognizing emotions, achieving high accuracy is still challenging due to the complex and non-stationary nature of these signals. A key challenge is extracting features that encapsulate different information aspects including temporal, frequency, and spatial information. This article presents a systematic review and analysis of the most significant machine and deep learning techniques for the EEG-based emotion recognition task, and outlines the challenges faced by the current models.

Keywords: EEG, Emotion Recognition, Subject-dependent, Subject-independent, Machine Learning, Deep Learning.

1 | Introduction

Emotions play a crucial part in day-to-day living as they can influence people's articulation, impressions, and decisions. Due to the quick growth of artificial intelligence, affective computing has emerged as a hot area of research for the human-computer interface, with applications spanning from driver fatigue detection [1] and autonomous control systems [2] to healthcare [3]. Most recent research on emotion recognition has centered on analyzing body language, voice patterns [4], and facial expression analysis [5]. However, these methods may face difficulties in accurately identifying emotional states, especially when people intentionally repress their emotional expressions. In light of limitations encountered with traditional emotion recognition methods, recent scientific investigations have pivoted towards the analysis of physiological signals by examining a range of intrinsic indicators, including skin temperature (SKT), heart rate (HR) [6], electrocardiography (ECG), electromyography (EMG), electrooculography (EOG), and electroencephalography (EEG) [7,8], this method aims to determine real emotional states.

EEG is a complex time-series data set that is utilized in the clinical and psychiatric sectors to detect brain activity as well as to record and store EEG signals. Particularly in the last 10 years, the discipline of brain-computer interface (BCI) has made use of EEG data to mimic human brain function. Automatic emotion recognition [9], is a crucial area of research in human-computer interaction, where we must recognize various emotional states using EEG signals. Two models—the discrete model and the dimensional model—are used

Corresponding Author: AMNasser@zu.edu.eg

to study and classify various emotional states [10]. Six basic emotional states are included in the discrete model: fear, anger, disgust, surprise, pleasure, and sadness. Emotional states are separated into two-dimensional areas known as valence and arousal by the dimensional model. The valence scale goes from depressed to joyful, and the arousal scale from bored to inspired. The interactions of valence and arousal spaces can be used to identify discrete emotional states [11]. The ideal BCI uses transitory EEG waves to detect the arousal or valence of an emotional state without explicit user input [12].

Different tasks are investigated in EEG emotion recognition including subject-dependent and subject-independent tasks. Subject-dependent tasks tailor EEG emotion recognition to individual users by using the same subject's data for training and testing. Conversely, subject-independent models are designed to generalize across various subjects and aim to develop robust classifiers that perform well on new subjects without needing personalized calibration [12].

Several deep learning and machine learning techniques are utilized to address EEG recognition of emotional challenges. These methods typically consist of three stages: pre-processing, feature extraction, and classification. Before further analysis, the pre-processing step cleans the data from noise and artifacts. The feature extraction step is the main stage before implementing machine learning or deep learning models. Due to the complexity and nonlinear nature of EEG signals, effectively leveraging their characteristics to obtain more complementary and discriminative data representations is a challenge in EEG-based emotion recognition. Many existing methods in EEG signal analysis focus solely on temporal, spectral, or spatial aspects, often failing to comprehensively capture all three. Temporal analysis examines changes in signals over time. Spectral analysis assesses the power distribution across various frequency bands, while spatial features refer to the spatial distribution of EEG signals across different EEG channels.

Moreover, approaches leveraging deep learning to enhance signal representation face limitations due to the high computational demands of such models [9]. Additionally, several methods employ window segmentations, with or without overlap, to increase the dataset size by treating these segments as distinct EEG signals with the same emotional state as the original. These segments can also be fed into a deep model to capture temporal data. However, these approaches often lose long-term dependencies, and overlapping can result in redundant information [13–15].

This work explores various machine and deep learning techniques used in EEG-based emotion recognition. It focuses on studies from the past five years and includes insights from relevant survey papers to provide a well-rounded overview of the field. The paper is organized as follows: **Section 2** offers a detailed background and context for understanding the progress and applications in emotion recognition. **Section 3** reviews recent research on machine and deep learning techniques, grouped by subject-dependent and subject-independent tasks, while also discussing the key challenges faced in these studies. Finally, **Section 4** summarizes the main conclusions of this research.

2 | Emotion Recognition Background

Emotion recognition is a rapidly advancing field that intersects various disciplines such as psychology, neuroscience, and computer science. The ability to accurately identify and understand human emotions has profound implications for enhancing human-computer interactions, improving mental health diagnostics, and developing responsive systems in various applications, from entertainment to healthcare. In this chapter, we will discuss the fundamental concepts of emotions, the methods used to analyze and classify emotional states, and the techniques employed in emotion recognition. We will delve into the discrete and dimensional methods of categorizing emotions, explore the role of EEG in emotion recognition, and examine various EEG-based emotion recognition tasks and approaches. This comprehensive overview will provide the necessary background and context for understanding the advancements and applications in the emotion recognition field.

2.1 |Emotions

Because humans are social creatures by nature, they rely heavily on alliances and partnerships of all kinds to survive in a world where resources are scarce. The ability to distinguish or comprehend Emotions that are significant to society in others and in ourselves, as well as to use them to guide and direct our social interactions, is essential to the development of these social partnerships. An individual's feelings and cognitive states are primarily correlated with their innate biological states, which are known as emotions. In essence, emotions are an internal assessment and response to perceived events resulting from interactions with different external stimuli [16–18]. Therefore, noticeable behavioral alterations caused by the emotional states encountered include changes to verbal intonations, body language, and facial expressions [19,20].

Studies and discoveries about emotions have captured people's interest for many decades. According to the most recent WHO figures, 970 million individuals worldwide had mental disorders related to emotions in 2019. Depression and Anxiety were the most common disorders, with increases of 26% and 28% in just one year [21]. Considering the significance of studying and recognizing emotions, two representation methods are employed to analyze and classify various emotional states: the discrete method and the dimensional method [22–24].

2.1.1| Discrete Representation Method

The discrete method comprises the following six basic emotional states: fear, anger, disgust, surprise, joy, and sadness [25]. Robert Plutchik expanded on this with his wheel of emotions, which includes eight primary emotions: fear, joy, sadness, trust, anger, surprise, anticipation, and disgust as shown in Figure 1 [26].

2.1.2| Dimensional Representation Method

The dimensional method offers a framework for recognizing the complexity and richness of emotional states, enabling a more in-depth analysis by dividing emotional states into dimensional spaces, typically two-dimensional (2D) and three-dimensional (3D) emotional space models [22–25].

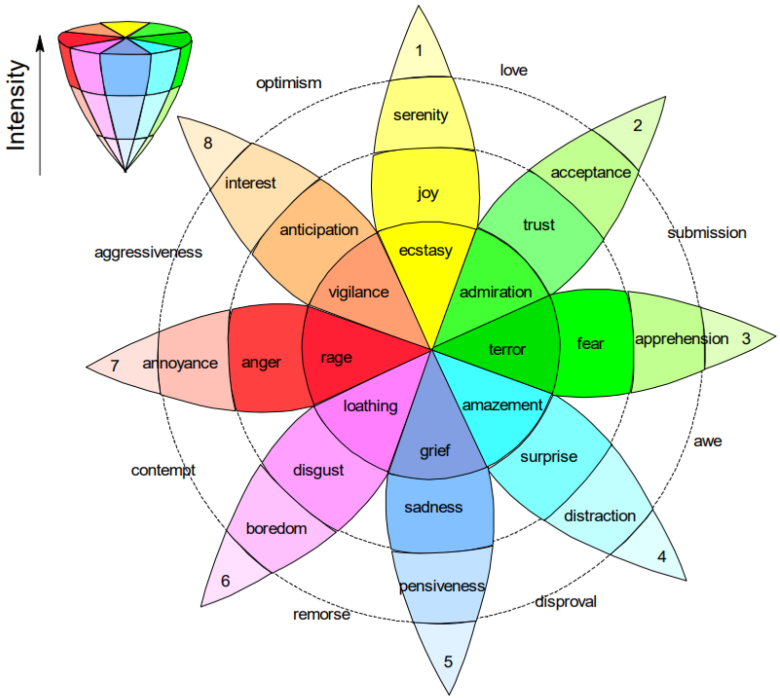

Figure 1. Plutchik's Wheel of Emotions [26].

In the 2D emotional space model, emotions are classified along two axes: valence (V), which ranges from positive (Pos) to negative (Neg), and arousal (A), which indicates the level of activation from high to low. The ratings or values on the valence axis, which range from Pos to Neg, represent an individual's level of happiness or sadness. Similar to this, an activated state (excitement) is indicated by a high value on the arousal axis, whereas an inactivated state (calmness) is shown by a low value. The combination of valence and arousal spaces (V-A spaces) can be used to identify discrete emotional states as shown in Figure 2 [25,27].

Similarly, the 3D emotional space model expands on this by incorporating an additional dimension: dominance (D). The individual's level of dominant control over their emotional process is reflected in the dominant dimension. A user's emotional state (such as surprise, fear, etc.) is at a lower dominance level while they are under the influence of their surroundings. On the other hand, a user's emotional state reaches a greater dominance degree when they can manage their external surroundings. The 3D model maps emotions along the axes of valence-arousal-dominance spaces (V-A-D spaces), providing a more comprehensive representation of emotional states as shown in Figure 3 [28,29].

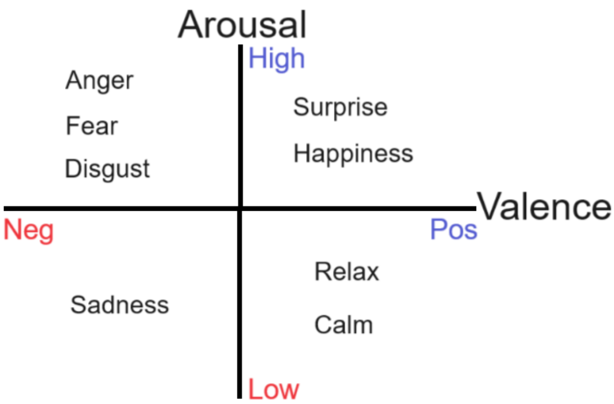

Figure 2. Valence-Arousal scale [25].

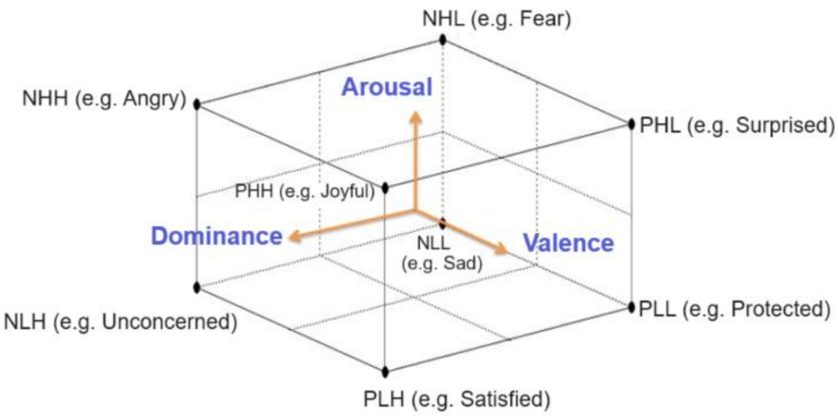

Figure 3. Valence-Arousal-Dominance space – N stands for negative; P stands for positive; H stands for high; L stands for low. The order of notions is mapped respectively to Valence, Arousal, and Dominance. E.g. PHH means positive valence-high arousal-high dominance [29].

It is essential to note that there is a one-to-one relationship between the discrete emotional states and particular places in the continuous dimensional state space. For example, the emotions of sadness and happiness are positioned in negative valence-low Arousal-low Dominance and positive valence-high Arousal-high Dominance respectively.

2.2 |Emotion Recognition

Emotion, a complex psychological state, manifests itself in both physical behaviors and physiological activities [30]. Emotion recognition involves identifying an individual's emotional state through their actions or physiological responses. For effective emotion research, quickly and accurately determining people's emotional states is essential to achieving a more natural and enhanced human-computer interaction.

2.2.1| Emotion Recognition Techniques

Over the last decade, significant efforts have been devoted to recognizing emotions by analyzing affective information obtained from various sources, including voices, neurophysiological devices' signals, videos, and texts [19,20]. The core of emotion recognition models lies in employing statistical ML techniques including classification, clustering, or regression) to classify human emotions in real-time or offline. Therefore, this task is particularly a challenge due to the need to extract and leverage latent components within noisy data from multiple sources such as body language, text, eye look, voice, facial emotions, bodily movements, and bio-signals.

Currently, assessing emotional states derived from physiological processes is a prominent topic in Affective Computing. Various psychophysiological studies have indicated dependencies between physiological processes and emotion cognition, although the exact order of these processes remains debated [31]. Some commonly utilized physiological modalities for detecting emotional responses include EEG, electrocardiogram (ECG), electrodermal activity, and electrooculogram [32–34]. Given that the autonomic nervous system is regulated and controlled by the central nervous system (the brain) during emotional processes, it is especially beneficial for research to directly use brain activity data (such as EEG) to evaluate emotional states and understand the mechanisms behind emotional cognition.

2.2.2| Potential Research and Applications of Emotion Recognition

Emotion recognition, especially EEG-based, offers a wide range of potential applications. For instance, the development of emotion-aware driver assistance technologies for vehicles is considered a promising way to improve driving safety [35]. In neurology, analyzing emotions and corresponding neural activities in response to specific stimuli can aid in diagnosing affective disorders such as post-traumatic stress disorder [36] and depression [37]. In the leisure and entertainment sector, particularly computer gaming, Researchers seek to identify players' emotional states to modify the degree of challenge, penalty, and encouragement in the game [38]. In Virtual Reality (VR) applications, in education, for instance, Emotional states have been shown to affect memory, and positive feelings have been found to improve spatial learning. Therefore, recognizing emotions during learning in a VR environment is crucial [39].

2.3 |EEG

EEG signal is a random signal with discrete spatial characteristics and non-stationary behavior. It uniquely captures changes in scalp potential directly. Unlike other physiological signals, EEG provides a more genuine and reliable reflection of the human condition, particularly in a resting state. Due to continuous advancements in brain science and signal processing technology, EEG signals have become increasingly popular in the emotion recognition field. The EEG is a method for recording electrophysiological signals indicative of brain activity. It represents the summation of extracellular field potentials generated by postsynaptic potentials due to the activation of many neurons in the brain [40].

Although the process that generates the EEG signal is a complicated one, it nonetheless contains a lot of useful information. This signal is time-varying, spatially discrete, and non-stationary. Its waveform is typified by an absence of well-defined rules. It is therefore challenging to summarize its behavior. But from a frequency domain standpoint, the EEG output appears to be rhythmic, with distinct rhythms exhibiting distinct meanings and also called EEG signal frequency bands [40,41]. EEG frequency bands are distinct ranges of frequencies within the EEG signal, each associated with specific mental or physiological states as described in Table 1. These frequency bands help researchers and clinicians understand different aspects of brain activity and their correlation with various mental states and neurological functions.

2.3.1| EEG Signals Acquisitions

One important step in the procedure is the placement of electrodes on the scalp to acquire EEG data. There are two primary classifications for this technique: invasive and non-invasive. Although invasive procedures are less noisy and have better accuracy, their usage is restricted because of safety concerns. Based on the tools utilized, non-invasive techniques can be further divided into two categories: dry and wet electrodes. To minimize resistance interference and improve signal stability, wet electrodes need a conductive medium to be placed between them and the scalp [42,43], as shown in Figure 4. However, wet electrode equipment has a shorter lifespan due to dielectric wear and can be uncomfortable for subjects due to the thick electrolytes applied to the scalp. Dry electrodes, while avoiding the need for a conductive medium and enhancing subject comfort, are less sensitive to electrode-scalp contact, leading to higher interference and weaker signal strength. This can complicate feature extraction in subsequent analysis. Both types have their pros and cons, so the choice should depend on the experiment's duration and the specific EEG signal requirements.

Table 1. EEG signal frequency bands.

Frequency band	Frequency range	Associated states
Delta (δ)	1-4 HZ	Deep inactivity, dysfunctional states
Theta (θ)	4-7 HZ	Drowsy
Alpha (α)	7-15 HZ	Idle, rest
Beta (β)	15-30 HZ	Mixed
Gamma (γ)	30-300 HZ	Activation
Sigma (σ)	300 HZ-1 KHZ	Spiking activity

Figure 4. EEG data acquisition using a non-invasive method [43]

The equipment used to acquire EEG data is fitted with multiple electrodes designed to capture electrical signals from the scalp-brain interface. The number of electrodes used can vary depending on the device, including 16, 32, and 64 electrodes. These electrodes are positioned on the scalp in a strategic manner using accepted placement techniques such as the 10–5, 10–10, and 10–20 systems, where these values represent the original space between the skulls in percentage [44]. These methods assist in gathering EEG readings from various brain areas. The most popular electrode placement location for the 10–20 system is illustrated in Figure 5. The proportion of the entire length of the skull that separates the distance between neighboring electrodes is represented by the numbers 10 and 20.

Each electrode location is labeled with a letter indicating the lobe and a number indicating the hemisphere location. F stands for Frontal, T for Temporal, C for Central, O for Occipital, and P for Parietal. The letter z (zero) indicates a midline electrode placement. The right hemisphere is represented by an even number, while the left hemisphere is represented by odd values [44]. Figure 6 shows the brain lobes and Table 2 shows the electrodes placement position of the 10/20 standard system on the different brain lobes.

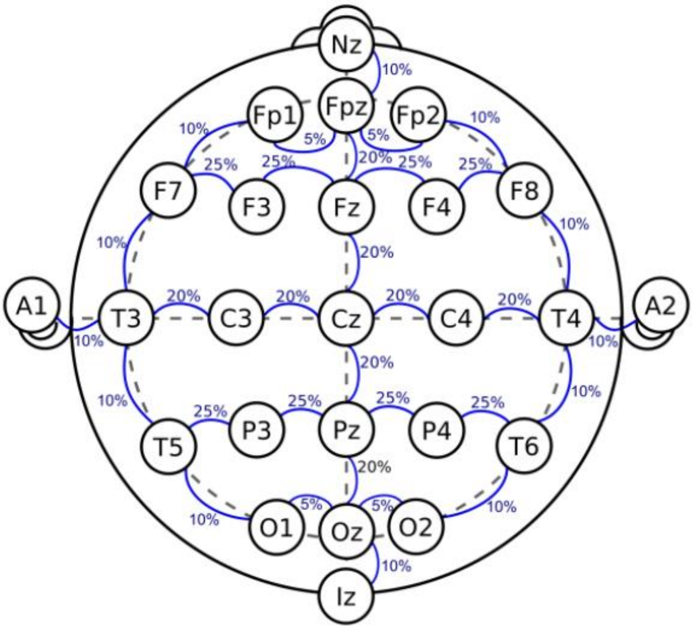

Figure 5. Standard 10/20 system of electrode placement method [44].

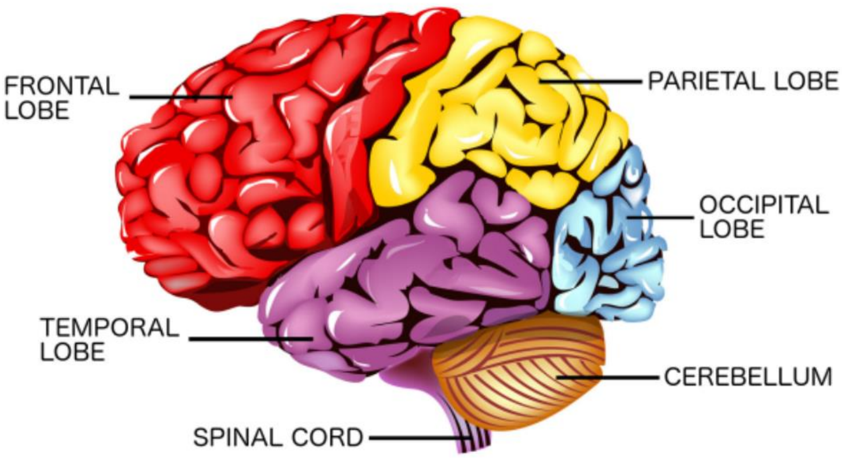

Figure 6. Brain lobes [44]

Table 2. Electrodes placement position of 10/20 system on the different brain lobes [44] bands.

Brain lobe	Hemisphere	Locations of electrodes
F (frontal)	Left	Fp1, AF3, AF7, F1, F3, F5, F7, FC1, FC3, FC5
	Right	Fp2, AF4, AF8, F2, F4, F6, FC2, FC4, FC6
P (parietal)	Left	C1, C3, C5, CP1, CP5, P1, P3, P5
	Right	C2, C4, CP2, CP4, CP6, P2, P4, P6
T (temporal)	Left	FT7, T3, T5, TP7
	Right	FT8, T4, T6, TP8
O (occipital)	Left	O1, PO3, PO7
	Right	O2, PO4, PO8

2.4 | EEG-based Emotion Recognition Tasks

Different tasks are investigated in EEG emotion recognition tasks including subject-dependent and subject-independent tasks. Subject-dependent models tailor EEG emotion recognition to individual users by the same subject's training and testing data. However, they lack generalizability across different subjects, making them impractical for large-scale applications. Conversely, subject-independent models are designed to generalize across diverse users by training on a dataset with EEG signals from various subjects. These models aim to develop robust classifiers that perform well on new subjects without needing personalized calibration.

2.5 | EEG-based Emotion Recognition Approaches

EEG emotion recognition task is addressed using a variety of deep learning and machine learning techniques. Pre-processing, feature extraction, and classification are the three stages that these methods typically include. To prepare the data for further analysis, the pre-processing step eliminates any noise and artifacts. Before using machine learning or deep learning models, the feature extraction step is essential. Because of the strong non-stationary characteristics of EEG data, the extraction of such features is a difficult procedure requiring specialized understanding. Recent studies used multivariate statistical analysis techniques to capture hand-crafted features in the frequency, time-frequency, and nonlinear domains that may represent EEG characteristics [45]. For example, some trials were constructed to preserve the channels' spatial information based on differential entropy (DE) from building a two-D mapping representation which was fed into a convolutional neural network (CNN) compared with other machine learning techniques such as Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) [46]. In addition, frequency bands energy and Entropy are extracted and fed into KNN and SVM classifiers in [13] and integrated with a discriminative graph regularized extreme learning machine (GELM) in [47].

Recently, there has been an increase in the use of deep learning-based methods for enhancing the efficiency and applicability of EEG applications. To understand more about the relationships between different brain areas, conditional entropy and power spectral density (PSD) as frequency and non-linear features with CNN are proposed in [48]. In [49], the authors merged CNN and sparse autoencoder with time-domain characteristics such as statistical features, sample entropy, and Pearson correlation coefficient (PCC). In [9], different statistical characteristics are extracted from EEG signal frequency bands and integrated with a deep belief-conditional random field framework. Moreover, a custom CNN model is presented to capture asymmetric, regional, and temporal features [50]. In [51], another deep model is utilized including long-short-term memory neural networks (LSTM) and graph convolutional neural networks (GCNN), where these models work with extracted DE features. To address the limits of EEG data and improve the performance of emotional recognition, more enhancements based on generative models and data augmentation approaches were explored [52]. To count the minimum number of samples needed for calibration without adjusting pre-trained classification methods for every user, a few-shot learning methodology is introduced [53].

In conclusion, while significant progress has been made in EEG emotion recognition, ongoing research is needed to overcome the challenges of data quality, feature extraction, model generalization, and computational complexity. Continued advancements in both machine learning and deep learning techniques hold promise for further improving the accuracy and applicability of EEG-based emotion recognition systems. In the next section, the latest research in subject-dependent and subject-independent tasks are discussed, and the limitations and research gaps are identified.

3 |EEG Emotion Recognition Literature Review

Several models and features have been thoroughly studied and have contributed significantly to the field of EEG emotion recognition for subject-dependent and independent problems. EEG emotion recognition subject-dependent models are customized for each user and need to be trained and calibrated for every new subject. Despite its effectiveness, this method is constrained by its inability to be applied to various people. On the other hand, subject-independent models offer more adaptability and usefulness in real-world circumstances as they function over a wide range of subjects without requiring individual calibration. This section reviews the previous studies proposed to solve the EEG emotion recognition problem with subject-dependent and subject-independent tasks, highlighting the limitations and challenges in these models.

3.1 | Subject-dependent Task Previous Studies

In the rapidly expanding field of emotional computing, a great deal of research has been done on the categorization of emotions in subject-dependent tasks. Discrete wavelet transform (DWT) was applied to frontal channels in [13] to extract the energy and entropy of information from frequency bands after splitting the signal into 2-second long segments with 1-second overlapping. These features were then fed into SVM and KNN models. Using the DEAP dataset, this method produced accuracy values of 86.75% and 84.05% for two classes of V and A scales. The feature selection approach in [47] employed a minimal redundancy maximum relevance (MRMR) to extract six unique features, including PSD, DE, and asymmetry characteristics, from each window of the entire signal duration that lasted one second and included no overlap. After that, these characteristics were sent to discriminative GELM, which on the DEAP dataset produced four classes with an accuracy of 69.67%. The previous methods' disadvantage is that they apply static signal segmentation, which aims to capture temporal data but loses long-term dependencies and creates redundancy through overlapping. The DWT method captures temporal and spectral information but is more sensitive to time-frequency changes and not robust to signal noise. To address the interactions between different brain areas, 0.5-second overlapping frames were used in [48] to extract conditional entropy and PSD, which passed to the CNN as a heat map image to capture complement features. These features were then transferred to an LSTM network, which on the DEAP dataset produced accuracy values of 71.09%, 72.58%, and 74.77 for two classes of V, A, and Liking (Li) scales. Additionally, the study in [49] merged statistical features, PCC enhanced with sparse autoencoder, and sample entropy with CNN model, while splitting the signal into 14 segments with 4s overlapping. On the DEAP dataset, this method produced accuracy values of 89.49% and 92.86% for two classes of V and A scales, respectively. Moreover, a custom CNN model is presented to capture regional, temporal, and asymmetric information, where the model splits the signal into 60 segments [50]. On the DEAP dataset, the model produced 96.65% and 97.11% accuracy for 2-classes of V and A scales. Moreover, for the 2 classes, LSTM and GCNN were utilized based on DE after splitting data into 10 segments (6s wide), achieving 90.45% and 90.60% for V and A scales [51]. The prior models used static signal segmentation and the computationally difficult deep approach.

Furthermore, the first and second-order differential was captured from 59 segments with 1s overlap. These features are enhanced with a stacked autoencoder and passed as input to a variety of classifiers, including RF, KNN, and decision trees. Using the DEAP dataset, the accuracy for the V scale was 62.63% across three classes [14]. With another 20 segments of the EEG signals, the PCC images were built and passed to the CNN model, where 78.22%, 74.92% on two classes, and 70.23%, 70.25% on three classes were achieved by utilizing the DEAP dataset [54]. More advanced investigations were presented based on CNN with a self-attention mechanism, where CNN was built based on DE [55]. On the DEAP dataset, this model produced 95.15%, 95.76%, 95.64%, and 95.86% for 2 classes with V, A, D, and Li scales. In [56], a novel dilated CNN model was presented to work with 2D and 3D representation features captured from channels' spatial features and frequency bands, achieving 89.67% and 90.93% for 2 classes. To integrate wavelet transforms and CNN with DEAP dataset, Continuous wavelet transform (CWT) is utilized to capture scalograms, achieving 83.43% and 84.2616% for 4 and 8 classes [57]. Similar to the DWT, the CWT extracts temporal and spectral information, but it is less robust to signal noise and more sensitive to variations in time frequency.

Despite the wide range of research that has been done, the majority of emotion identification methods ignore separating the temporal and frequency dependence of the EEG data. Furthermore, several models that have been presented make use of brief seconds-long EEG signal segments. Specifically, a trial signal of duration (t) is split into (n) segments, each of which has a tiny (t/n) second long with an overlap or non-overlap sliding window, and each segment is assigned the same label representing the entire experience. These segmentation strategies may be incorrect since the subjects' emotional states may be continuously changing, and the reported emotion is for the complete simulation. Furthermore, the feature representation that is currently in use is not robust enough against noise, deformations, or changes in time shift. To improve emotion classification models, it is crucial to find stable EEG features that capture both temporal and frequency information for the entire signal. This has led to a more thorough analysis and an accurate depiction of the EEG signal.

3.2 | Subject-independent Task Previous Studies

The subject-independent proposed methods aim to generate solutions for emotion recognition that may be applied more broadly and accessed by a broader range of people [58]. In the work in [59], spectral entropy values were extracted from each EEG channel's data and employed as input in the bidirectional long-short-term memory (BI-LSTM) model, achieving an accuracy of 76.93% on the GAMEEMO dataset with 10-fold cross-validation (CV). In [60], each channel's holographic representation is created using features gathered from 5 sub-bands. The obtained features are then fed into a CNN as the feature extraction layer, where they are combined with SVM to identify two V and A classes. They achieved an accuracy of 76.61% and 77.72% for V and A using the DEAP dataset and a 10-fold CV. In [15], the EEG signal is divided into ten 6 s segments. Each segment's Hjorth parameters and wavelet (entropy, energy) characteristics are computed and concatenated to each other, and then emotions are identified using SVM. They achieved 75.19% and 81.74% for V and A of the 2-classes classification task and 73.62% for the 4-classes classification task with the LOSO validation technique on the DEAP dataset. This method's disadvantage is that it ignores temporal information in favor of frequency components. Its accuracy was less than 90%. They employed the computationally complex deep method.

A feature extraction approach based on the domain adaptive symmetric and positive definite (SPD) matrix network has been developed [61]. This approach integrated prototype learning with the Riemannian metric to compute the sample point diffusion matrix. They achieved 67.99% and 76.57% accuracy for V and A on the DEAP dataset with the LOSO validation technique. The disadvantage of their method is that they ignore both frequency and temporal information. They capture geometric and structural information, including relationships between channels and features, but achieve less than 90% accuracy. In [62], a hand-crafted framework based on a prime pattern network and tunable q-factor wavelet transform (TQWT) was introduced, integrated with SVM as a classification layer, achieving superior results over 99% on the DEAP and the GAMEEMO datasets. The usage of TQWT effectively captures frequency and temporal aspects but it is noise-sensitive, like other wavelet methods. Additionally, it overlooks inter-channel relationships by modeling each channel separately.

Moreover, the work in [63] captured a modified version of DE, power spectral density, and Hjorth parameters from 4 sub-bands then fed the result into BI-LSTM to learn temporal and spatial features and finally apply a fully connected layer to classify emotions. These features were created after dividing the EEG signals into 1s segments with 50% overlap. They achieved 73.50% and 75% for V and A on the DEAP dataset. In [64], the DE was captured from each channel after splitting the signal into 3 s segments long to build a feature pyramid network as a feature extraction layer with the SVM classification layers with an accuracy of over 80% on the DEAP dataset. Furthermore, in [65], the signal was split into 4 s segments with 2s overlaps to be used with the BI-LSTM network, achieving an accuracy of 61.39% and 63.35% on V and A on the DEAP dataset. The previous method's disadvantage is that it applies static signal segmentation, which aims to capture temporal data but loses long-term dependencies and creates redundancy through overlapping.

In [66], an approach to EEG-based emotion recognition was presented, which used multi-head attention mechanisms and graph convolution networks to model spatial brain activity. They achieved an accuracy of over 69% on the DEAP dataset. Their technique resulted in a lower accuracy rate and used a computationally complex deep method. A transfer learning framework based on the Google-Net network was presented in [67], where EEG signals were converted to EEG images using CWT, which is more sensitive to time-frequency changes and not robust to signal noise. They achieved an accuracy of 93.31% on the GAMEEMO with a 10-fold CV. In [68], the EEG waves were decomposed by Empirical Mode Decomposition (EMD) and other statistical features and then fed into Bi-LSTM, achieving an accuracy of 90.33% for 10-fold CV. The EMD can capture time and frequency information but is more sensitive to time-frequency changes.

In [69], an alternative method was proposed in which the EEG signal was separated into segments of 1s size, and the root mean square of each window was computed for 5 sub-bands then the data was input into a projection dictionary pair learning PDPL model for classification, achieving 49.01% accuracy on the GAMEMEO dataset. The disadvantage of their method is the usage of static signal segmentation and their accuracy is less than 90%.

In [70], Multivariate Fast Iterative Filtering (MFIF) was applied to the EEG signal to decompose it to a simpler form and Hjorth parameters with nine entropy features were extracted from each 2s segment with KNN as a classification layer. They achieved an accuracy of 97.857% with the GAMEMEO dataset with a LOSO CV and 98.45% with a 10-fold CV. Furthermore, a framework for the temporal-spectral recognition of emotions is suggested, in which high-level spectral characteristics were extracted using the Manifold-Net network and then fed into the BI-LSTM network for classification [71]. They achieved 71.88% and 69.38% accuracy for V and A on the DEAP dataset. The MFIF captured temporal and frequency aspects well but struggled with noise, like other wavelet methods. In [72], an information potential index was applied - after decomposing the EEG signal into five frequency bands - between EEG channels to find the spatial pattern. This study employed a KNN classifier to categorize emotions into four classes on the DEAP dataset and achieved results of 94.75% with the LOSO validation technique. They capture frequency and spatial information but ignore temporal information.

Despite the variety of existing studies, many existing methods in EEG signal analysis concentrate either on temporal, or spectral aspects, often neglecting to capture both comprehensively. Integrated approaches provide deeper insights into the signal characteristics. A few models comprehensively integrate spatial aspects alongside temporal and spectral dimensions. Moreover, those models fall short in accuracy and rely on computationally intensive deep-learning architectures, particularly with noise-sensitive and non-invariant methods. Moreover, approaches leveraging deep learning to enhance signal representation face limitations due to the high computational demands of such models. In addition, current models often limit EEG signal classification to a narrow range of emotional states or datasets, raising questions about their robustness and adaptability to a broader spectrum of emotions.

3.3 | Limitations and Challenges Summary

In conclusion, EEG-based emotion categorization techniques have evolved significantly, yet they encounter specific limitations and challenges:

- Many existing methods in EEG signal analysis concentrate either on temporal or spectral aspects, often neglecting to capture both comprehensively. Temporal analysis focuses on signal changes over time, while spectral analysis looks at power distribution across various frequency bands. Integrated approaches provide deeper insights into the signal characteristics.

- However, even when both dimensions are considered, the techniques may fall short in accuracy, particularly with noise-sensitive and non-invariant methods. Moreover, approaches leveraging deep learning to enhance signal representation face limitations due to the high computational demands of such models.
- A few models comprehensively integrate spatial aspects alongside temporal and spectral dimensions. Moreover, those models fall short in accuracy and rely on computationally intensive deep-learning architectures.
- Traditional static segmentation, with or without overlap, captured temporal data but lost long-term dependencies, while overlapping led to redundant information.
- Current models often limit EEG signal classification to a narrow range of emotional states or datasets, raising questions about their robustness and adaptability to a broader spectrum of emotions.
- The dataset limitation challenges.

Therefore, as a future direction, this paper suggests the following points:

- Explore other effective feature extraction techniques that are stable enough against deformation noise and variations due to the complexity of the EEG signals. In addition, it possesses strong discriminative ability, is computationally efficient, and contributes additional, distinct information within the selected subset of extracted features.
- Build the model in a way that captures different information aspects including temporal, spectral, and spatial information while working on the whole signal rather than applying static segmentation. This can lead to a more comprehensive analysis and a more accurate representation of EEG signals which in turn enhance emotion classification models.
- Regarding the dataset limitation problem, the development of novel data augmentation techniques customized to the properties of EEG signals may greatly enhance the efficacy of emotion recognition models.
- Investigating new deep learning architectures is crucial to enhancing EEG data-based emotion recognition performance.

4 | Conclusion

This study presents a systematic review and analysis of the most significant machine and deep learning techniques for the EEG-based emotion recognition task in two directions including subject-dependent and subject-independent tasks. In addition, it outlines limitations and challenges faced by the current models and finally suggests future directions.

Despite the variety of existing studies, many existing methods in EEG signal analysis concentrate either on temporal, or spectral aspects, often neglecting to capture both comprehensively. Integrated approaches provide deeper insights into the signal characteristics. A few models comprehensively integrate spatial aspects alongside temporal and spectral dimensions. Moreover, those models fall short in accuracy and rely on computationally intensive deep-learning architectures, particularly with noise-sensitive and non-invariant methods. In addition to the traditional static segmentation, with or without overlap, which captures temporal data but suffers from losing long-term dependencies, overlapping leads to redundant information.

Future work should focus on exploring new effective feature extraction techniques that are stable enough against deformation noise, and variations due to the complexity of the EEG signals. In addition to possessing strong discriminative ability, be computationally efficient, and contribute additional, distinct information within the selected subset of extracted features. In addition to extracting different aspects from the EEG including temporal, spectral, and spatial information, explore new deep learning architectures, and data augmentation techniques to enhance EEG data-based emotion recognition performance.

Acknowledgments

The author is grateful to the editor and reviewer, who offered assistance in the form of advice, assessment, and checking.

Funding

This research has no funding source.

Data Availability

The datasets generated during the current study are publicly available.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

References

[1] T. Chen, S. Ju, F. Ren, M. Fan, Y. Gu, EEG emotion recognition model based on the LIBSVM classifier, Measurement 164 (2020) 108047.

[2] L. Jiang, Y. Wang, Z. Tang, Y. Miao, S. Chen, Casting defect detection in X-ray images using convolutional neural networks and attention-guided data augmentation, Measurement 170 (2021) 108736.

[3] Y. Lu, M. Wang, W. Wu, Y. Han, Q. Zhang, S. Chen, Dynamic entropy-based pattern learning to identify emotions from EEG signals across individuals, Measurement 150 (2020) 107003.

[4] W. Li, Z. Zhang, A. Song, Physiological-signal-based emotion recognition: An odyssey from methodology to philosophy, Measurement 172 (2021) 108747.

[5] P. Surcinelli, F. Andrei, O. Montebarocci, S. Grandi, Emotion recognition of facial expressions presented in profile, Psychol Rep 125 (2022) 2623–2635.

[6] W. Wen, G. Liu, N. Cheng, J. Wei, P. Shangguan, W. Huang, Emotion Recognition Based on Multi-Variant Correlation of Physiological Signals, IEEE Trans Affect Comput 5 (2014) 126–140. https://api.semanticscholar.org/CorpusID:18814389.

[7] G. Keren, T. Kirschstein, E. Marchi, F. Ringeval, B. Schuller, End-to-end learning for dimensional emotion recognition from physiological signals, in: 2017 IEEE International Conference on Multimedia and Expo (ICME), 2017: pp. 985–990. https://doi.org/10.1109/ICME.2017.8019533.

[8] Y. Yang, Q. Gao, X. Song, Y. Song, Z. Mao, J. Liu, Facial Expression and EEG Fusion for Investigating Continuous Emotions of Deaf Subjects, IEEE Sens J 21 (2021) 16894–16903. https://doi.org/10.1109/JSEN.2021.3078087.

[9] H. Chao, Y. Liu, Emotion Recognition From Multi-Channel EEG Signals by Exploiting the Deep Belief-Conditional Random Field Framework, IEEE Access 8 (2020) 33002–33012. https://doi.org/10.1109/access.2020.2974009.

[10] E.L. van den Broek, Ubiquitous emotion-aware computing, Pers Ubiquitous Comput 17 (2011) 53–67. https://doi.org/10.1007/s00779-011-0479-9.

[11] P. Ekman, An argument for basic emotions, Cogn Emot 6 (1992) 169–200. https://doi.org/10.1080/02699939208411068.

[12] S.K. Khare, V. Blanes-Vidal, E.S. Nadimi, U.R. Acharya, Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations, Information Fusion 102 (2024) 102019. https://doi.org/10.1016/j.inffus.2023.102019.

[13] Z. Mohammadi, J. Frounchi, M. Amiri, Wavelet-based emotion recognition system using EEG signal, Neural Comput Appl 28 (2016) 1985–1990. https://doi.org/10.1007/s00521-015-2149-8.

[14] C. Qing, R. Qiao, X. Xu, Y. Cheng, Interpretable Emotion Recognition Using EEG Signals, IEEE Access 7 (2019) 94160–94170. https://doi.org/10.1109/access.2019.2928691.

[15] M. Khateeb, S.M. Anwar, M. Alnowami, Multi-Domain Feature Fusion for Emotion Classification Using DEAP Dataset, IEEE Access 9 (2021) 12134–12142. https://doi.org/10.1109/ACCESS.2021.3051281.

[16] T. Dalgleish, The emotional brain, Nat Rev Neurosci 5 (2004) 583–589.
[17] L.F. Barrett, Emotions are real., Emotion 12 (2012) 413.
[18] J.J. Gross, Emotion regulation: Current status and future prospects, Psychol Inq 26 (2015) 1–26.
[19] S. Planalp, Varieties of cues to emotion in naturally occurring situations, Cogn Emot 10 (1996) 137–154.
[20] K.R. Scherer, Vocal communication of emotion: A review of research paradigms, Speech Commun 40 (2003) 227–256.
[21] World Health Organization. Mental Disorders, Https://Www.Who.Int/News-Room/Fact-Sheets/Detail/Mental-Disorders/?Gad_source=1&gclid=Cj0KCQjwk6SwBhDPARIsAJ59GwfPLhgdxZysvhRHEbockNo7g8vHt3Y3TnHmhfEoJedNmyuLWbMre5AaAg-_EALw_wcB (2022).
[22] C.E. Izard, Basic emotions, relations among emotions, and emotion-cognition relations., (1992).
[23] J. Posner, J.A. Russell, B.S. Peterson, The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology, Dev Psychopathol 17 (2005) 715–734.
[24] R.W. Levenson, Basic emotion questions, Emotion Review 3 (2011) 379–386.
[25] J.A. Russell, A circumplex model of affect., J Pers Soc Psychol 39 (1980) 1161.
[26] A. Petrauskas, Plutchik's Wheel and The Unity of Opposites, (n.d.).
[27] A. Raheel, S.M. Anwar, M. Majid, Emotion recognition in response to traditional and tactile enhanced multimedia using electroencephalography, Multimed Tools Appl 78 (2018) 13971–13985. https://doi.org/10.1007/s11042-018-6907-3.
[28] Y. Zhao, D. Xie, R. Zhou, N. Wang, B. Yang, Evaluating users' emotional experience in Mobile libraries: An emotional model based on the pleasure-arousal-dominance emotion model and the Five factor model, Front Psychol 13 (2022) 942198.
[29] A. Mehrabian, Framework for a comprehensive description and measurement of emotional states., Genet Soc Gen Psychol Monogr 121 (1995) 339–361.
[30] R. Adolphs, D.J. Anderson, The neuroscience of emotion: A new synthesis, (2018).
[31] W.B. Cannon, The James-Lange theory of emotions: A critical examination and an alternative theory, Am J Psychol 39 (1927) 106–124.
[32] A. Dzedzickis, A. Kaklauskas, V. Bucinskas, Human emotion recognition: Review of sensors and methods, Sensors 20 (2020) 592.
[33] Y.R. Veeranki, H. Kumar, N. Ganapathy, B. Natarajan, R. Swaminathan, A systematic review of sensing and differentiating dichotomous emotional states using audio-visual stimuli, IEEE Access 9 (2021) 124434–124451.
[34] A. Horvers, N. Tombeng, T. Bosse, A.W. Lazonder, I. Molenaar, Detecting emotions through electrodermal activity in learning contexts: A systematic review, Sensors 21 (2021) 7869.
[35] Z. Halim, M. Rehan, On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning, Information Fusion 53 (2020) 66–79.
[36] V. Rozgic, A. Vazquez-Reina, M. Crystal, A. Srivastava, V. Tan, C. Berka, Multi-modal prediction of ptsd and stress indicators, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2014: pp. 3636–3640.
[37] H. Cai, Z. Qu, Z. Li, Y. Zhang, X. Hu, B. Hu, Feature-level fusion approaches based on multimodal EEG data for depression recognition, Information Fusion 59 (2020) 127–138.
[38] T.B. Alakus, M. Gonen, I. Turkoglu, Database for an emotion recognition system based on EEG signals and various computer games – GAMEEMO, Biomed Signal Process Control 60 (2020) 101951. https://doi.org/https://doi.org/10.1016/j.bspc.2020.101951.
[39] K.E. Stavroulia, M. Christofi, E. Baka, D. Michael-Grigoriou, N. Magnenat-Thalmann, A. Lanitis, Assessing the emotional impact of virtual reality-based teacher training, The International Journal of Information and Learning Technology 36 (2019) 192–217.
[40] F.L. DA SILVA, Niedermeyer's electroencephalography: basic principles, clinical applications, and related fields, (2010).
[41] W. Wang, G. Zhang, L. Yang, V.S. Balaji, V. Elamaran, N. Arunkumar, Revisiting signal processing with spectrogram analysis on EEG, ECG and speech signals, Future Generation Computer Systems 98 (2019) 227–232.
[42] N. Panigrahi, S.P. Mohanty, Brain Computer Interface: EEG Signal Processing, CRC Press, 2022.
[43] H.A. Hamzah, K.K. Abdalla, EEG-based emotion recognition systems; comprehensive study, Heliyon 10 (2024) e31485. https://doi.org/10.1016/j.heliyon.2024.e31485.
[44] S.M. Alarcao, M.J. Fonseca, Emotions recognition using EEG signals: A survey, IEEE Trans Affect Comput 10 (2017) 374–393.
[45] X. Li, Y. Zhang, P. Tiwari, D. Song, B. Hu, M. Yang, Z. Zhao, N. Kumar, P. Marttinen, EEG Based Emotion Recognition: A Tutorial and Review, ACM Comput Surv 55 (2022) 1–57. https://doi.org/10.1145/3524499.
[46] J. Li, Z. Zhang, H. He, Hierarchical Convolutional Neural Networks for EEG-Based Emotion Recognition, Cognit Comput 10 (2017) 368–380. https://doi.org/10.1007/s12559-017-9533-x.
[47] W.-L. Zheng, J.-Y. Zhu, B.-L. Lu, Identifying Stable Patterns over Time for Emotion Recognition from EEG, IEEE Trans Affect Comput 10 (2019) 417–429. https://doi.org/10.1109/taffc.2017.2712143.
[48] Siddharth, T.-P. Jung, T.J. Sejnowski, Utilizing Deep Learning Towards Multi-Modal Bio-Sensing and Vision-Based Affective Computing, IEEE Trans Affect Comput 13 (2022) 96–107. https://doi.org/10.1109/taffc.2019.2916015.
[49] J. Liu, G. Wu, Y. Luo, S. Qiu, S. Yang, W. Li, Y. Bi, EEG-Based Emotion Classification Using a Deep Neural Network and Sparse Autoencoder, Front Syst Neurosci 14 (2020) 43. https://doi.org/10.3389/fnsys.2020.00043.

[50] H. Cui, A. Liu, X. Zhang, X. Chen, K. Wang, X. Chen, EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network, Knowl Based Syst 205 (2020) 106243. https://doi.org/10.1016/j.knosys.2020.106243.

[51] Y. Yin, X. Zheng, B. Hu, Y. Zhang, X. Cui, EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM, Appl Soft Comput 100 (2021) 106954. https://doi.org/10.1016/j.asoc.2020.106954.

[52] Y. Luo, L.-Z. Zhu, Z.-Y. Wan, B.-L. Lu, Data augmentation for enhancing EEG-based emotion recognition with deep generative models, J Neural Eng 17 (2020) 056021. https://doi.org/10.1088/1741-2552/abb580.

[53] S. Bhosale, R. Chakraborty, S.K. Kopparapu, Calibration free meta learning based approach for subject independent EEG emotion recognition, Biomed Signal Process Control 72 (2022) 103289. https://doi.org/10.1016/j.bspc.2021.103289.

[54] Md.R. Islam, Md.M. Islam, Md.M. Rahman, C. Mondal, S.K. Singha, M. Ahmad, A. Awal, Md.S. Islam, M.A. Moni, EEG Channel Correlation Based Model for Emotion Recognition, Comput Biol Med 136 (2021) 104757. https://doi.org/10.1016/j.compbiomed.2021.104757.

[55] D. Li, L. Xie, B. Chai, Z. Wang, H. Yang, Spatial-frequency convolutional self-attention network for EEG emotion recognition, Appl Soft Comput 122 (2022) 108740. https://doi.org/10.1016/j.asoc.2022.108740.

[56] Q. Yao, H. Gu, S. Wang, X. Li, A Feature-Fused Convolutional Neural Network for Emotion Recognition From Multichannel EEG Signals, IEEE Sens J 22 (2022) 11954–11964. https://doi.org/10.1109/jsen.2022.3172133.

[57] O. Almanza-Conejo, D.L. Almanza-Ojeda, J.L. Contreras-Hernandez, M.A. Ibarra-Manzano, Emotion recognition in EEG signals using the continuous wavelet transform and CNNs, Neural Comput Appl 35 (2022) 1409–1422. https://doi.org/10.1007/s00521-022-07843-9.

[58] D.W. Prabowo, H.A. Nugroho, N.A. Setiawan, J. Debayle, A systematic literature review of emotion recognition using EEG signals, Cogn Syst Res 82 (2023) 101152. https://doi.org/https://doi.org/10.1016/j.cogsys.2023.101152.

[59] T.B. Alakus, I. Turkoglu, Emotion recognition with deep learning using GAMEEMO data set, Electron Lett 56 (2020) 1364–1367. https://doi.org/https://doi.org/10.1049/el.2020.2460.

[60] A. Topic, M. Russo, Emotion recognition based on EEG feature maps through deep learning network, Engineering Science and Technology, an International Journal 24 (2021) 1442–1454. https://doi.org/https://doi.org/10.1016/j.jestch.2021.03.012.

[61] Y. Wang, S. Qiu, X. Ma, H. He, A prototype-based SPD matrix network for domain adaptation EEG emotion recognition, Pattern Recognit 110 (2021) 107626. https://doi.org/https://doi.org/10.1016/j.patcog.2020.107626.

[62] A. Dogan, M. Akay, P.D. Barua, M. Baygin, S. Dogan, T. Tuncer, A.H. Dogru, U.R. Acharya, PrimePatNet87: Prime pattern and tunable q-factor wavelet transform techniques for automated accurate EEG emotion recognition, Comput Biol Med 138 (2021) 104867. https://doi.org/10.1016/j.compbiomed.2021.104867.

[63] V.M. Joshi, R.B. Ghongade, IDEA: Intellect database for emotion analysis using EEG signal, Journal of King Saud University - Computer and Information Sciences 34 (2022) 4433–4447. https://doi.org/https://doi.org/10.1016/j.jksuci.2020.10.007.

[64] F. Hou, Q. Gao, Y. Song, Z. Wang, Z. Bai, Y. Yang, Z. Tian, Deep feature pyramid network for EEG emotion recognition, Measurement 201 (2022) 111724. https://doi.org/https://doi.org/10.1016/j.measurement.2022.111724.

[65] M. Zhu, Q. Wang, J. Luo, Emotion Recognition Based on Dynamic Energy Features Using a Bi-LSTM Network, Front Comput Neurosci 15 (2022). https://www.frontiersin.org/articles/10.3389/fncom.2021.741086.

[66] D. Priyasad, T. Fernando, S. Denman, S. Sridharan, C. Fookes, Affect recognition from scalp-EEG using channel-wise encoder networks coupled with geometric deep learning and multi-channel feature fusion, Knowl Based Syst 250 (2022) 109038. https://doi.org/https://doi.org/10.1016/j.knosys.2022.109038.

[67] M. Aslan, CNN based efficient approach for emotion recognition, Journal of King Saud University - Computer and Information Sciences 34 (2022) 7335–7346. https://doi.org/https://doi.org/10.1016/j.jksuci.2021.08.021.

[68] A. Abdulrahman, M. Baykara, T.B. Alakus, A Novel Approach for Emotion Recognition Based on EEG Signal Using Deep Learning, Applied Sciences 12 (2022). https://doi.org/10.3390/app121910028.

[69] J. Su, J. Zhu, T. Song, H. Chang, Subject-Independent EEG Emotion Recognition Based on Genetically Optimized Projection Dictionary Pair Learning, Brain Sci 13 (2023). https://doi.org/10.3390/brainsci13070977.

[70] R. Vempati, L.D. Sharma, EEG rhythm based emotion recognition using multivariate decomposition and ensemble machine learning classifier, J Neurosci Methods 393 (2023) 109879. https://doi.org/https://doi.org/10.1016/j.jneumeth.2023.109879.

[71] R. Li, C. Ren, S. Zhang, Y. Yang, Q. Zhao, K. Hou, W. Yuan, X. Zhang, B. Hu, STSNet: a novel spatio-temporal-spectral network for subject-independent EEG-based emotion recognition, Health Inf Sci Syst 11 (2023) 25. https://doi.org/10.1007/s13755-023-00226-x.

[72] A. Goshvarpour, A. Goshvarpour, EEG emotion recognition based on an innovative information potential index, Cogn Neurodyn (2024). https://doi.org/10.1007/s11571-024-10077-1.